

- Proposition (inclusion-exclusion identity): If $A_1, A_2, \dots, A_n$ are any n events, let



$$\sigma_1 = \sum_{i=1}^n P(A_i),$$

$$\sigma_2 = \sum_{1 \leq i < j \leq n} P(A_i \cap A_j),$$

$$\sigma_3 = \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k),$$

$$\dots = \dots$$

$$\sigma_k = \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k})$$

$$\dots = \dots$$

$$\sigma_n = P(A_1 \cap A_2 \cap \dots \cap A_n).$$

then

$$P(A_1 \cup \dots \cup A_n) = \sigma_1 - \sigma_2 + \sigma_3 - \dots + (-1)^{k+1} \sigma_k + \dots + (-1)^{n+1} \sigma_n.$$

generalization of ④

Q: For an outcome w contained in m out of the n events, how many times is its probability $p(w)$ repetitively counted in $\sigma_1, \dots, \sigma_n$?

An intuitive proof for discrete Ω :
For $w \in \Omega$,

$p(w)$ counted m times in σ_1

$= \dots = \binom{m}{2} = \dots = \sigma_2$

$\vdots$

$\vdots = \binom{m}{k} = \dots = \sigma_k$

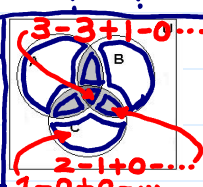
$\vdots$

$\vdots = \binom{m}{m} = \dots = \sigma_m$

$\vdots$

$\vdots = 0 = \dots = \sigma_{m+1}$

$$[-\binom{m}{0}] + \binom{m}{1} - \binom{m}{2} + \binom{m}{3} - \dots + (-1)^{k+1} \binom{m}{k} + \dots + (-1)^{m+1} \binom{m}{m} = 0 \quad (\text{LNp.2-7})$$



proof. Prove by induction.

$n=2$, it holds

by ④, LNp.3-11

$$(A_1 \cap A_2) \cup (A_2 \cap A_3)$$

$$n=3, P(A_1 \cup A_2 \cup A_3) = P(A_1 \cup A_2) + P(A_3) - P((A_1 \cup A_2) \cap A_3)$$

$$= P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

by ④

$$= P(A_1 \cap A_3) + P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3)$$

$$= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3)$$

For $n > 3$, by mathematical induction (exercise)

Notes:

- There are $\binom{n}{k}$ summands in σ_k

- In symmetric examples,

symmetric outcomes (LNp.3-4)

cf.

$$\sigma_k = \binom{n}{k} P(A_1 \cap \dots \cap A_k)$$

- It can be shown that

for proof. See textbook. p.44

$$P(A_1 \cup \dots \cup A_n) \leq \sigma_1$$

$$P(A_1 \cup \dots \cup A_n) \geq \sigma_1 - \sigma_2$$

$$P(A_1 \cup \dots \cup A_n) \leq \sigma_1 - \sigma_2 + \sigma_3$$

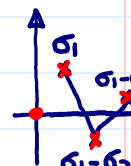
...

3rd proposition in LNp.3-11

σ_1 加太多

$\sigma_1 - \sigma_2$ 減太多

$\sigma_1 - \sigma_2 + \sigma_3$ 加太多



item → Example (The Matching Problem). *classical prob.* = $\#A/\#\Omega$

Symmetric outcomes

- Applications: (a) Taste Testing. (b) Gift Exchange.

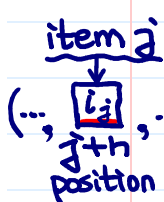
- Let Ω be all permutations $\omega = (i_1, \dots, i_n)$ of $1, 2, \dots, n$.

Thus, $\#\Omega = n!$.

eg. (3, 1, 5, ...)
 1st 2nd 3rd

none of them get his/her own gift

at least one person get his/her own gift



Let

event of interest can be difficult to identify

$$A_j = \{\omega: i_j = j\} \text{ and } A = \bigcup_{j=1}^n A_j$$

Q: $P(A) = ?$ (What would you expect when n is large?)

- By symmetry,

$$\sigma_k = \binom{n}{k} P(A_1 \cap \dots \cap A_k),$$

$P(A) \uparrow ?$ to 1?

$P(A) \downarrow ?$ to 0? or others?

- Here,

$$P(A_1) = \frac{1 \times (n-1)!}{n!} = \frac{1}{n},$$

$$P(A_1 \cap A_2) = \frac{(n-2)!}{n!} = \frac{1}{(n)_2},$$

$$\dots = \dots,$$

$$P(A_1 \cap \dots \cap A_k) = \frac{(n-k)!}{n!} = \frac{1}{(n)_k}, \dots$$

for $k = 1, \dots, n$.



- So,

$$\sigma_k = \binom{n}{k} \frac{1}{(n)_k} = \frac{1}{k!}, = \frac{n!}{k! (n-k)!} \times \frac{(n-k)!}{n!}$$

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$$

$$P(A) = \sigma_1 - \sigma_2 + \dots + (-1)^{n+1} \sigma_n = \sum_{k=1}^n (-1)^{k+1} \frac{1}{k!},$$

$$P(A) = 1 - \sum_{k=0}^{\infty} (-1)^k \frac{1}{k!} \approx 1 - \frac{1}{e} = 0.632 \Rightarrow P(A^c) \approx e^{-1} = 0.368$$

when $n \rightarrow \infty$

- Note: approximation accurate to 3 decimal places if $n \geq 6$.

Proposition: If $A_1, A_2, \dots$ is a partition of Ω , i.e.,

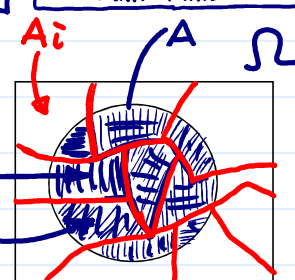
- $\bigcup_{i=1}^{\infty} A_i = \Omega$,
- $A_1, A_2, \dots$ are mutually exclusive,

then, for any event $A \subset \Omega$,

$$P(A) = \sum_{i=1}^{\infty} P(A \cap A_i).$$

proof. $A = A \cap \Omega = A \cap (\bigcup_{i=1}^{\infty} A_i) = \bigcup_{i=1}^{\infty} (A \cap A_i)$

$$P(A) = P(\bigcup_{i=1}^{\infty} (A \cap A_i)) \stackrel{\text{mutually exclusive}}{=} \sum_{i=1}^{\infty} P(A \cap A_i)$$



mutually exclusive

additivity axiom

Probability Measure for Continuous Sample Space

Q: How to define probability in a continuous sample space?

cf. → How to define P.M. for discrete Ω

uncountably infinite Ω

- Monotone Sequences of sets (LNp.3-7) → check Examples in LNp.3-1 3-6

Definition: A sequence of events $A_1, A_2, \dots$, is called increasing if

$$A_1 \subset A_2 \subset \dots \subset A_n \subset A_{n+1} \subset \dots \subset \Omega$$

and decreasing if

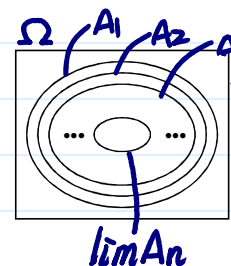
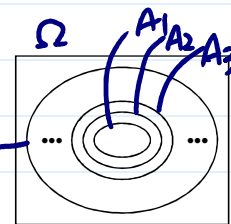
$$A_1 \supset A_2 \supset \dots \supset A_n \supset A_{n+1} \supset \dots \supset \emptyset$$

The limit of an increasing sequence is defined as

$$\lim_{n \rightarrow \infty} A_n = \bigcup_{i=1}^{\infty} A_i$$

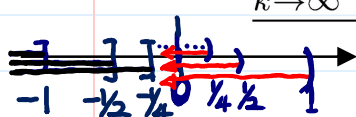
and the limit of an decreasing sequence is

$$\lim_{n \rightarrow \infty} A_n = \bigcap_{i=1}^{\infty} A_i$$



Example: If $\Omega = \mathbb{R}$ and $A_k = (-\infty, 1/k)$, then A_k 's are decreasing and

$$\lim_{k \rightarrow \infty} A_k = \{\omega : \omega < 1/k \text{ for all } k \in \mathbb{Z}_+\} = (-\infty, 0]$$



(exercise) $A_k = (-\infty, -1/k] \uparrow, \lim_{k \rightarrow \infty} A_k = (-\infty, 0)$

Proposition: If $A_1, A_2, \dots$, is increasing or decreasing, then

LNp.3-3
DeMorgan's Law

$$\left(\lim_{n \rightarrow \infty} A_n \right)^c = \lim_{n \rightarrow \infty} A_n^c$$

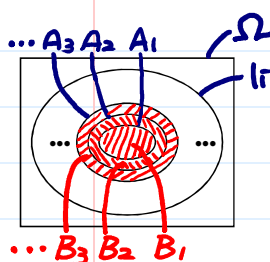
$$\left(\bigcup_{n=1}^{\infty} A_n \right)^c = \bigcap_{n=1}^{\infty} A_n^c = \lim_{n \rightarrow \infty} A_n^c$$

$$\left(\bigcap_{n=1}^{\infty} A_n \right)^c = \bigcup_{n=1}^{\infty} A_n^c = \lim_{n \rightarrow \infty} A_n^c$$

Proposition: If $A_1, A_2, \dots$, is increasing or decreasing, then

$$\lim_{n \rightarrow \infty} P(A_n) = P\left(\lim_{n \rightarrow \infty} A_n\right) \quad \text{⑤}$$

What A_n 's should be defined first?



proof. ① A_n increasing.

$$\text{Let } B_1 = A_1, B_2 = A_2 \cap A_1^c, \dots, B_n = A_n \cap A_{n-1}^c, \dots$$

$$\Rightarrow \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n, \bigcap_{n=1}^{\infty} B_n = \bigcap_{n=1}^{\infty} A_n = A_k, B_i \cap B_j = \emptyset \text{ for } i \neq j.$$

$$P\left(\lim_{n \rightarrow \infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} P(B_n) = \lim_{K \rightarrow \infty} \sum_{n=1}^K P(B_n)$$

$$= \lim_{K \rightarrow \infty} P\left(\bigcup_{n=1}^K B_n\right) = \lim_{K \rightarrow \infty} P\left(\bigcup_{n=1}^K A_n\right) = \lim_{K \rightarrow \infty} P(A_K)$$

by Ax3
Additivity

② A_n decreasing $\Rightarrow A_n^c$ increasing

$$1 - P\left(\lim_{n \rightarrow \infty} A_n\right) = P\left(\left(\lim_{n \rightarrow \infty} A_n\right)^c\right)$$

$$= P\left(\lim_{n \rightarrow \infty} A_n^c\right) \stackrel{\text{by ①}}{=} \lim_{n \rightarrow \infty} P(A_n^c) = \lim_{n \rightarrow \infty} 1 - P(A_n) = 1 - \lim_{n \rightarrow \infty} P(A_n)$$

countable sums

Example (Uniform Spinner): Let $\Omega = (-\pi, \pi]$. Define

Actually, it's enough to define P on any $(a, b]$, where a, b are rational numbers. (∵ rational numbers is a dense set)

$$P((a, b]) = \frac{b - a}{2\pi}$$

continuous sample space

play a similar role like $\omega_1, \omega_2, \dots$, in small p for discrete sample space (c.f. how to define P in LNp.3-7)

for subintervals $(a, b] \subset \Omega$. Then, extend P to other subsets using the 3 axioms. For example, if $-\pi < a < b < \pi$,

cumulative distribution function

$$P([a, b]) = P\left(\left(\bigcap_{k=1}^{\infty} \left(a - \frac{1}{k}, b\right]\right) \cap \Omega\right) = P\left(\bigcap_{k=1}^{\infty} \left(\left(a - \frac{1}{k}, b\right] \cap \Omega\right)\right)$$

Note: $[a, b] \neq (a, b]$

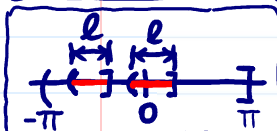
⇒ P not defined on $[a, b]$ under (*)

$$\lim_{k \rightarrow \infty} P\left(\left(a - \frac{1}{k}, b\right] \cap \Omega\right)$$

$$\lim_{k \rightarrow \infty} \frac{1}{2\pi} \left(b - a + \frac{1}{k}\right) = \frac{b - a}{2\pi}$$

$$A_k \downarrow \lim_{k \rightarrow \infty} A_k = [a, b]$$

基礎事件
discrete: $\{\omega_i\}$
continuous: $(a, b]$



Some notes

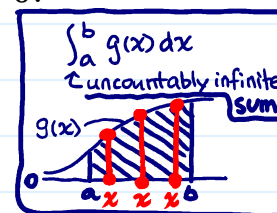
Note: $P([a, b])$ is derived through $P((a, b])$.

$P(\text{height} = 170) = 0?$

$$P(\{a\}) = P([a, b] - (a, b]) = P([a, b]) - P((a, b]) = 0.$$

$$\text{If } C = \{\omega_1, \omega_2, \dots\} \subset \Omega, \text{ then } \frac{b-a}{2\pi} \quad \frac{b-a}{2\pi}$$

$$P(C) = \sum_{i=1}^{\infty} P(\{\omega_i\}) = 0 + 0 + \dots = 0.$$



discrete sample space

In Ax3, countable unions (LNp.3-6)

The probability of all rational outcomes is zero

❖ Reading: textbook, Sec. 2.6

a countable (but dense) set.

Objective vs. Subjective "Interpretation" of Probability

- Evaluate the following statements

obj → 1. This is a fair coin → tossing the coin → repeatable case ← cf. → one-time case

subj → 2. It's 90% probable that Shakespeare actually wrote Hamlet

- Q: What do we mean if we say that the probability of rain tomorrow is 40%? → how to interpret it?

Objective: Long run relative frequencies

Subjective: Chosen to reflect opinion

$$P(A) = 0.4$$

$$O(A) = \frac{0.4}{0.6} = \frac{2}{3}$$

- The Objective (Frequency) Interpretation 頻率

Through Experiment: Imagine the experiment repeated N times. For an event A , let

$$N_A = \# \text{ occurrences of } A.$$

Then,

random number

$$P(A) \equiv \lim_{N \rightarrow \infty} \frac{N_A}{N}$$

frequency

By Law of Large Number (future lecture or course)

➤ Example (Coin Tossing):

	N	100	1000	10000	100000
random	N_H	55	493	5143	50329
frequency	N_H/N	.550 <i>.580</i>	.493 <i>.512</i>	.514 <i>.498</i>	.503 <i>.501</i>

→ 0.5

The result is consistent with $P(H)=0.5$.

★ odds of an event (勝敗率, 賠率)

• The Subjective Interpretation

$$o(A) = \frac{P(A)}{P(A^c)} = \frac{P(A)}{1-P(A)} \Rightarrow P(A) = \frac{o(A)}{1+o(A)}$$

➤ Strategy: Assess probabilities by imagining bets

➤ Example:

押 賺

2:1 odds on A

$$o(A) \in [0, \infty)$$

$$P(A) \uparrow \Leftrightarrow o(A) \uparrow \Leftrightarrow \alpha(A) \downarrow$$

$$(Peter) \text{ true } o(A) = \frac{P(A)}{P(A^c)} \geq \frac{2}{1}$$

$$1 - P(A)$$

event A

- Peter is willing to give two to one odds that it will rain tomorrow. His subjective probability for rain tomorrow is at least $2/3$

$$\Rightarrow P(A) \geq 2 - 2P(A)$$

$$\Rightarrow P(A) \geq 2/3$$

- Paul accepts the bet. His subjective probability for rain tomorrow is at most $2/3$

1:2 odds on A^c

$$(Paul) \text{ true } o(A^c) = \frac{P(A^c)}{P(A)} \geq \frac{1}{2}$$

$$\Rightarrow P(A) \leq 2/3$$

Bayesian approach

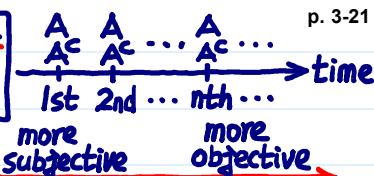
➤ Probabilities are simply personal measures of how likely we think it is that a certain event will occur

degree of personal belief

e.g. 2 in Lnp. 3-19.

➤ This can be applied even when the idea of repeated experiments is not feasible

repeatable case



p. 3-21

Summary

